Trends in the Presidential Management Fellows Program

Kyle Klarup
Last Updated: April 27, 2017

Table of Contents:

Introduction

The Presidential Management Fellows (PMF) Program bills itself as “a flagship leadership development program at the entry level for advanced degree candidates” [1]. Run out of the executive branch for over 30 years, anyone (well, mostly U.S. citizens) who will soon receive or recently received an advanced degree may apply to the program once a year during a 2-week window.

If accepted to the program, candidates don’t automatically get a position in the federal government. Instead, they become ‘Finalists’ and are given access to a special job board (the Talent Management System, or TMS). If they successfully apply to one of the opportunities posted on the TMS, they become

'Fellows,' are classified as appointed, and start a 2-year fellowship working for the federal government. More information on the PMF program and how to apply may be found here: at www.pmf.gov.

"Leadership development" could mean many things, so I wanted to know what sort of jobs would be available and which agencies were hiring before dedicating time to the application process. But I had to be a Finalist to view the job board. It was a lovely Catch-22. I applied anyway.

Soon after becoming a Finalist, I wrote some Python web scrapers and data analyzers to collect and process the information I wasn't able to access before. Partially, I did this for the practice and for my own enjoyment. But I hope the results might also help future prospective PMFs. This report contains wordy summaries of what I found, with the Python notebooks and much of the raw data available on GitHub at https://github.com/keklarup/PMF.

Trends over time in the PMF application process

Not all the information I wanted required becoming a Finalist and scraping the job board. Starting from at least 2012, the PMF Office has published yearly pdfs of the number of applicants, the number of selected Finalists, and the names and institutions of every Finalist.

Even before I applied to the PMF program, I looked at those pdfs to uncover any trends in the application process over the recent years. A word of caution on the results that follow: the structure of the PMF program is always changing. A separate track for STEM degree holders existed from 2014 through 2016 [2]. The 'semi-Finalist' status and in-person interview were removed in 2017 [3]. I don't have enough information to distinguish how these changes affected the following trends, and as the program continues to change, these trends might become completely irrelevant.

First, here are the applicant and Finalist numbers for the past 6 years.

Year	Applicants	Selected non-STEM Finalists	Selected STEM Finalists	Total Selected Finalists	Acceptance Rate
2012	9,100	628	N/A	628	6.9%
2013	12,120	663	N/A	663	5.5%
2014	7,000	518	91	609	8.7%
2015	7,800	508	92	600	7.7%
2016	6,050	455	97	552	9.1%
2017	6,370	417	N/A	417	6.5%

One of the first things to jump out from this data is a general drop in applicants from 2012-2016. It is likely that this trend isn't just statistical fluctuations, but it can't be ruled out. There was a slight rise in the number of applicants for 2017, but that number still fell well within the confidence interval of the 2012-2016 trajectory and didn't change the trend in any significant way.

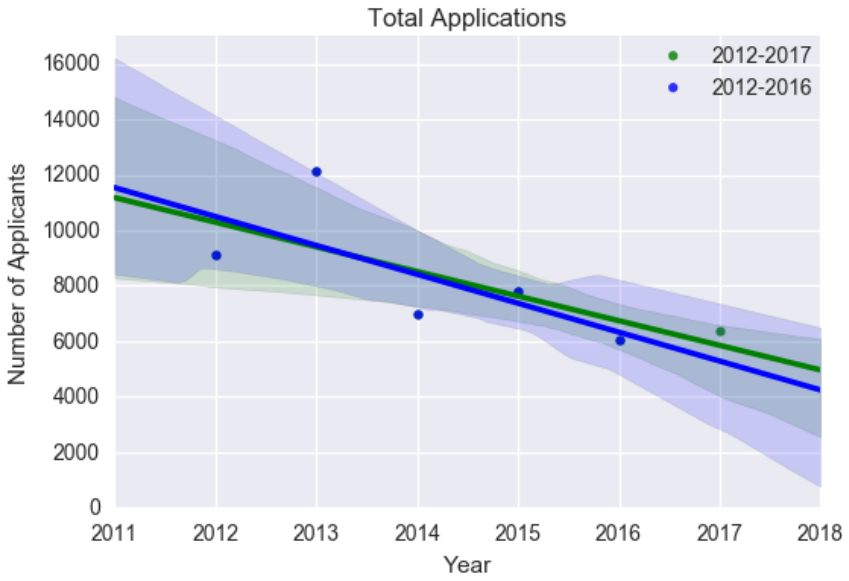


Figure 1: Linear fit trend lines for yearly applicant numbers, with shaded regions representing a 90% confidence interval. The best fit using records from 2012-2016 is shown in blue. The number of applicants for 2017 falls within the confidence window for the predicted trend and the inclusion of the 2017 recording does little to alter the fit (green line).

The number of selected Finalists also dropped over the same interval, but not at a more gradual pace. As a result, the acceptance rate began to creep up. However, 2017 saw a large drop in the number of accepted Finalists, which heavily slowed, if not completely stopped, that acceptance rate trend.

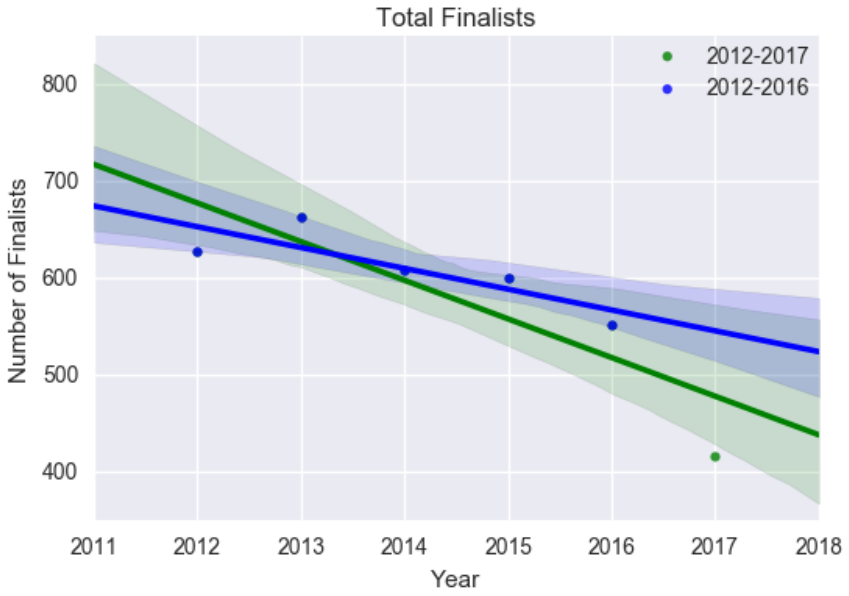


Figure 2: Linear fit trend lines for yearly Finalist numbers, with shaded regions representing a 90% confidence interval. The best fit using records from 2012-2016 is shown in blue. The number of chosen Finalists in 2017 falls well short of the predicted trend and the inclusion of the 2017 recording substantially alters the fit (green line).

In addition to these broad, macro numbers, the pdfs included the name of the degree granting institution for each Finalist. I could use these records to look for trends in which schools were sending

the most people into the PMF program. I became interested in this question because the PMF program has the goal of "developing a cadre of potential government leaders" [1]. There could be undesired effects if most of those potential leaders were coming from just a few institutional backgrounds.

From 2012-2017, 453 universities were represented in the 2,869 total Finalists. In a completely uniform world, every university would have sent about 1.5 Finalists every year. In reality, some schools are pipelines that send substantially more students into the program, while other schools often have years without any Finalists. The 2 questions I considered are:

1) How far away from a uniform distribution is reality, and
2) How is the real distribution changing with time.

To measure "far away" from a uniform distribution, I'm currently using a Kullback-Leibler divergence [3].

$$D_{KL}(P||Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)}$$

Essentially, this is a measurement of the difference between two probability distributions P and Q, where the larger the value of $D_{KL}(P||Q)$, the greater the divergence between the 2 distributions. However, this is my first time working with this particular metric, so I may have an incomplete grasp on the subtleties.

If I were looking for the best model to fit to the real distribution, I would set the real distribution of the finalists-to-schools as Q and compare that distribution to various models represented at P:

Distribution (n=453)	Uniform $x_i = c$	Linear $x_i = (n - i) + b$	Power $x_i = i^{-.9}$	Power Exponential $x_i = e^{i^{-.314}}$
$D_{KL}(P\|\|Q)$	.961	.516	.060	.015

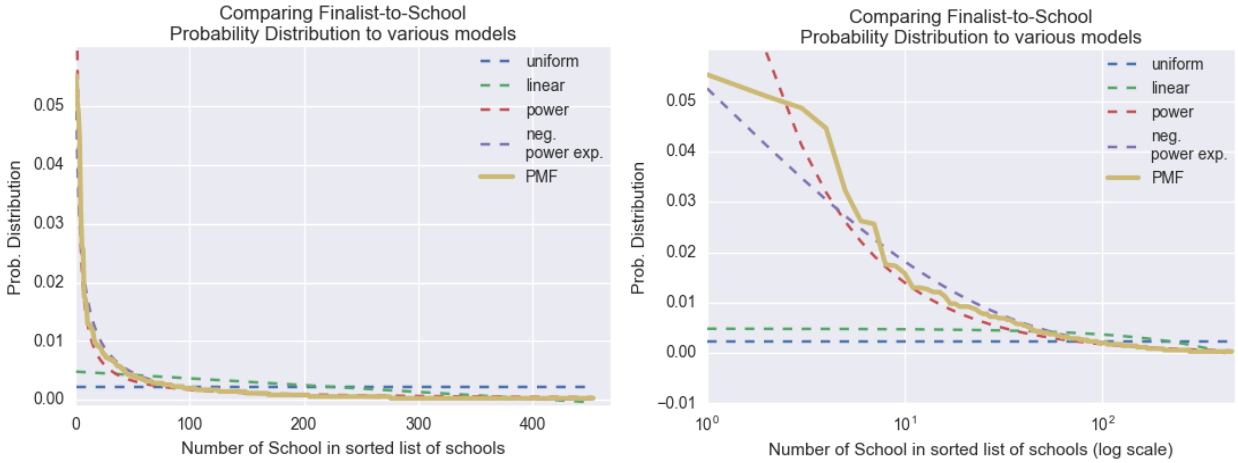


Figure 3: Examples of various types of fits to the real distribution of Finalists over the set of schools. The lowest Kullback-Leibler divergence value corresponds to the neg. power exponential fit. Left: Linear horizontal axis. Right: Log horizontal axis.

The lowest Kullback-Leibler divergence comes from a power-exponential fit, but I didn't bother doing an exhaustive search because I'm interested in the opposite question: How does the yearly real distribution compare to a uniform distribution?

To answer that, I'll switch up the terms in the Kullback-Leibler divergence definition. Now the uniform distribution will be Q and the yearly real distribution will be P.

Year	Applicants	Accepted Finalists	Acceptance Rate	$D_{KL}(P\|\|Q)$
2012	9100	628	6.9%	1.36
2013	12120	663	5.5%	1.11
2014	7000	609	8.7%	1.54
2015	7800	600	7.7%	1.53
2016	6050	552	9.1%	1.68
2017	6370	417	6.5%	1.82

The gradual increase in $D_{KL}(P||Q)$ over time indicates the real distribution of Finalists-to-schools is moving further away from uniform. It appears the cause of this increasing divergence is that the top Finalist producing schools (pipelines) are contributing a larger percentage of the total finalists at the expense of the rest of the institutions. This can be visualized in 2 ways.

First, plotting the yearly real distributions shows the probably a Finalist came from a pipeline school has increased over time:

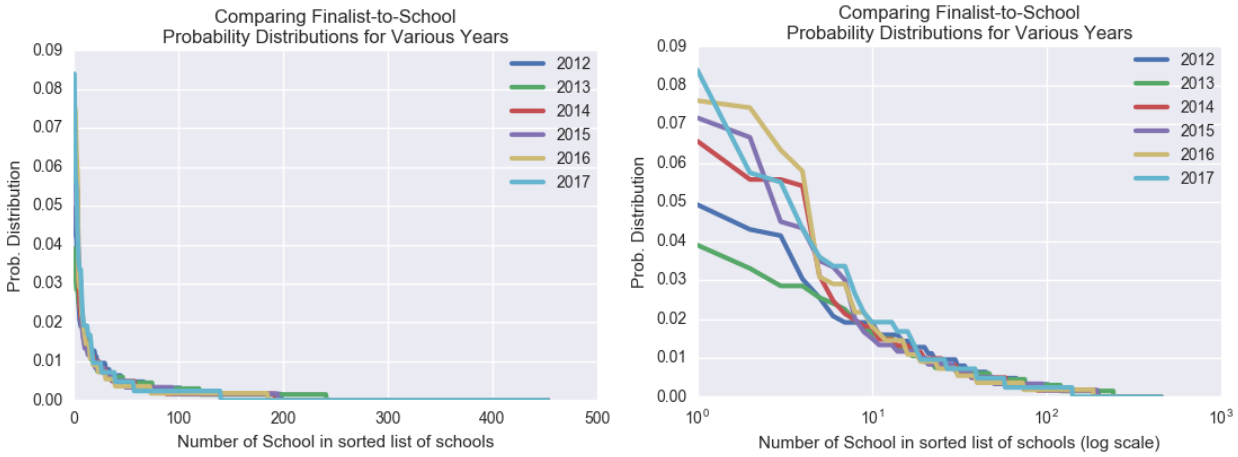


Figure 4: The real probability distributions of Finalists over the set of schools, sorted from the school that produced the most Finalists that year to the school that produced the fewest. The increasing probability for Finalists to come from schools that produce the greatest quantity of Finalists indicates an increase in the disparity between those schools and the rest of the set. Left: linear horizontal scale. Right: log horizontal scale.

Second, this trend in the dominance of pipeline schools can be seen by taking the top 5% most producing schools for every year and calculating the percent of total Finalists that come from those schools. In a uniform distribution, it should be 5%. If it's increasing, the top schools are producing a increasing share of the Finalists:

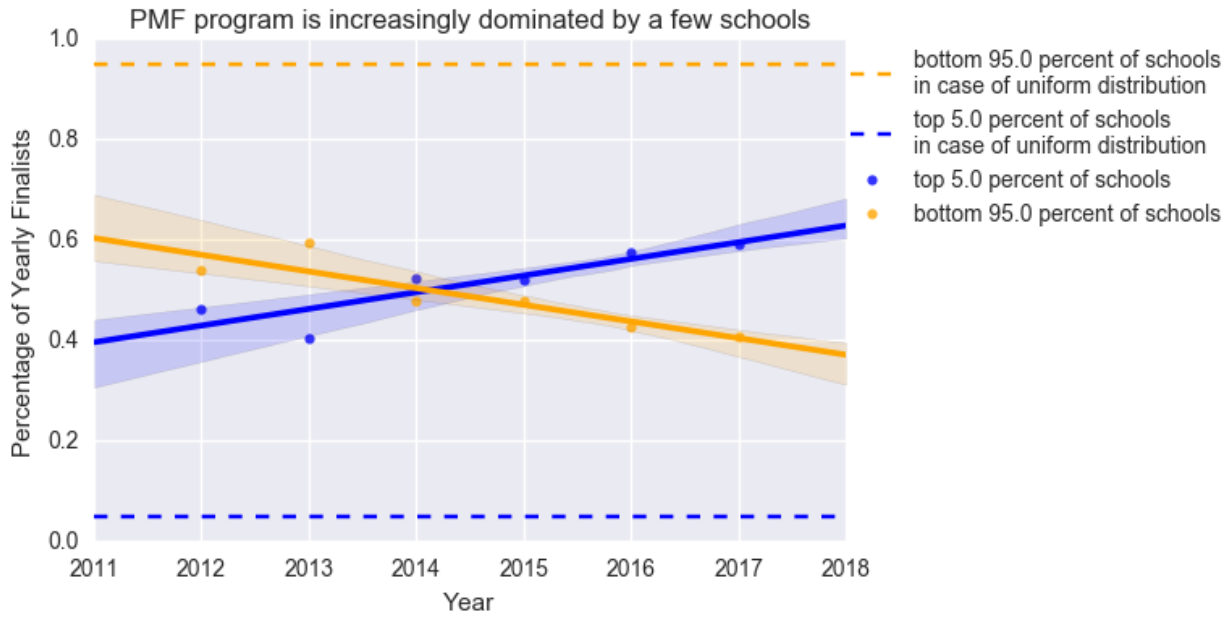


Figure 5: Best linear fit trendline for percent of PMF Finalists from the top 5% of schools (blue), as sorted by number of Finalists per school, vs the remaining 95% of schools (yellow). The shaded area represents a 90% confidence interval. Comparison curves for a uniform distribution of Finalists over the schools shown as dashed lines.

An equivalent observation is as the number of Finalists has dropped, the cuts have disproportionately affected the non-pipeline schools.

In an attempt to find a reason some schools have become pipelines for the PMF program, I have begun looking at the distance from each school to D.C. As the PMF application process pre-2017 required traveling to D.C. for an interview at the applicant's expense, geography might help account for the growth in pipelines. This work is not yet complete, but there are hints that distance to D.C. might be a contributing factor.

Here are the top 10 Finalists producing schools, accounting for over a third of all Finalists from 2012-2017, and the distance from each school to D.C:

University	Number of Finalists (2012-2017)	Distance to D.C. (miles)
GEORGETOWN UNIVERSITY	192	0
GEORGE WASHINGTON UNIVERSITY	177	0
AMERICAN UNIVERSITY	169	0
JOHNS HOPKINS UNIVERSITY	155	41
HARVARD UNIVERSITY	112	543
COLUMBIA UNIVERSITY	91	237
UNIVERSITY OF MICHIGAN - ANN ARBOR	89	525
GEORGE MASON UNIVERSITY	61	21
YALE UNIVERSITY	60	305
EMORY UNIVERSITY	55	632

Something that immediately jumps out about that list is proximity. The top 3 schools are in D.C, and the average distance to D.C. for the top 10 is around 200 miles.

It's possible those school have also contributed an increasing percentage of total applicants, and the trend in the Finalist purely reflects this distribution. It's also possible there is some selection bias during the admission process. I don't have the data to support any one theory over another

The application window for 2017 was open for 14 days (336 hours). With about 30 hours to go, I noticed the PMF Office was updating the application webpage with 3 numbers: the time left till the end of the application period, the number of started applications, and the number of finished applications. Each update wiped the previous numbers from the page. Still, I recorded 4 unique updates. I don't have any interesting insights from these numbers, besides the fact that many applicants waited till the final day to finish (or even start!) their application:

Hours to Deadline	Started Applications	Completed Applications	Percent Completed
336	0	0	0%
30	6500	2600	40.0%
11.83	7300	3600	49.3%
2.75	8000	5100	63.8%
1.5	8000	5600	70.0%
0	8000*	6370**	79.6%

* estimated. Assumed no applications started in last 90 minutes.
** from final official tally

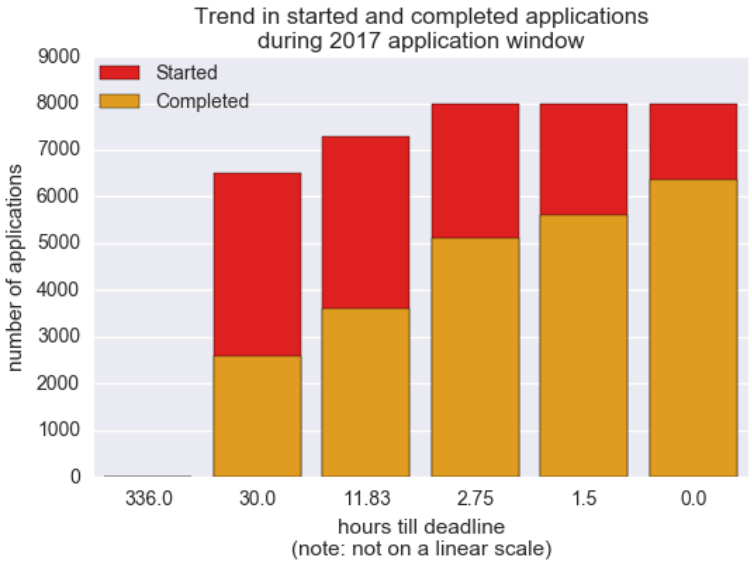


Figure 6: Recorded numbers of started and completed applications for the 2017 PMF class.

The 2012-2017 pdfs reporting the number of applicants, number of Finalists, and names and institutions of each applicant were found via a Google search of the pmf.gov website. The information from those reports was copied into a single csv file for doing analysis on the distribution of Finalists over the represented institutions. Complications arose during this analysis, because of two main issues: First, there was a large degree of variance in how institutions were named for each Finalist. Second, the process used to copy the pdf information into the csv file occasionally caused the name of a Finalist to be saved as part of the institution name. To deal with these problems, I wrote a semi-autonomous GUI to help me bin the institutional names.

Below is an example table of the variance and the relabeled institution names.

Year	**Type**	**School**	**School2**
2012	PMF	DUKE UNIV ? SCHOOL OF PUBLIC SERVICE	DUKE UNIVERSITY
2012	PMF	DUKE UNIV ? GRAD SCH ARTS & SCIENCE	DUKE UNIVERSITY
2016	PMF	DUKE UNIV/NICHOLAS SCH OF ENVIRON.	DUKE UNIVERSITY
2017	PMF	DUKE UNIVERSITY	DUKE UNIVERSITY
2015	PMF	KYLE DUKE UNIV/NICHOLAS SCH OF ENVIRON.	DUKE UNIVERSITY

A csv file for the 2012-2017 Finalists lists including year, track (if applicable), original institution name, and cleaned institution name for every Finalist is available on GitHub, as can the Python script used to create the GUI and do the pipeline analysis.

Resources	Filename	Description
1	PMF_Yearly_Finalists_Github.csv	Information on every Finalists from 2012-2017
2	Ktinker on Finalists-GitHub.ipynb	Python notebook to run re-naming GUI and do analysis on pipelines

Trends in Finalists

In addition to the yearly published pdfs, the PMF Office maintains an online list of current Finalists. This page contains more information than the pdfs, including the advanced degree of every Finalist and if they have been appointed. Because this list is only designed to show current Finalists, I only have these more detailed records for the 2016 and 2017 classes. Still, I'll include some information that can be gleaned from those records here.

First, the 2016 class fell in the typical appointment rate percentage window of 50-60%, at 57% [3]. I made 2 recordings of the 2016 class list. Once on January 12th, 2017 and again on March 2nd, 2017. The eligibility window for the 2016 class ended in late February, and from these 2 recordings I learning 7.4% of the 2016 appointments were recorded in the last 6 weeks of the 2016 Finalists' year of eligibility.

The 2017 class is just starting to see Finalists get appointed. But this time I'm monitoring the appointment rate in near real time and making many recordings of the number of appointments listed on the pmf.gov Finalist list. This information might be useful for estimating how long it takes to go from a tentative offer to an appointment, as well as letting current Finalists know how their appointment

experience compares to the class. As of this writing, 7.4% of 2017 Finalists had been reported as appointed.

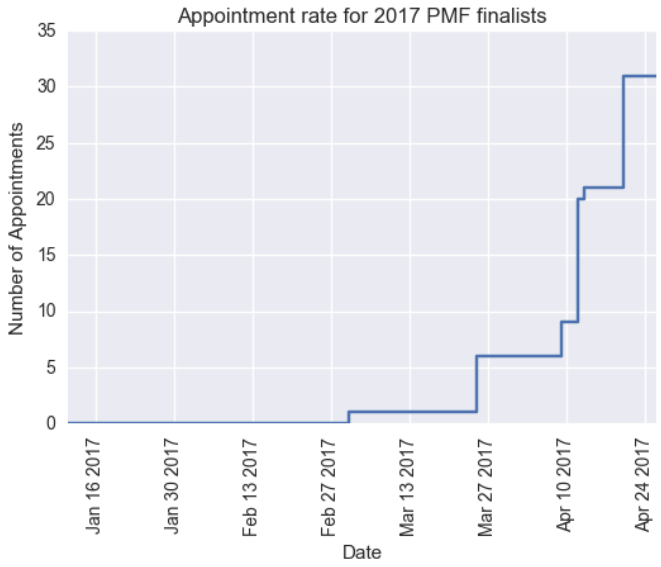


Figure 7: Recordings of the number of 2017 Finalists reported as Appointed on pmf.gov. Record from March 25, 2017 was lost and reconstructed from memory—might be off by a few days and/or an appointee.

There are 91 unique advanced degrees in the detailed lists of the 2016 and 2017 Finalists, although half of those degrees are held by 3 or fewer Finalists. I'll include a count plot of the number of Finalists for every degree in case future PMFs are curious how common their field is in the program.

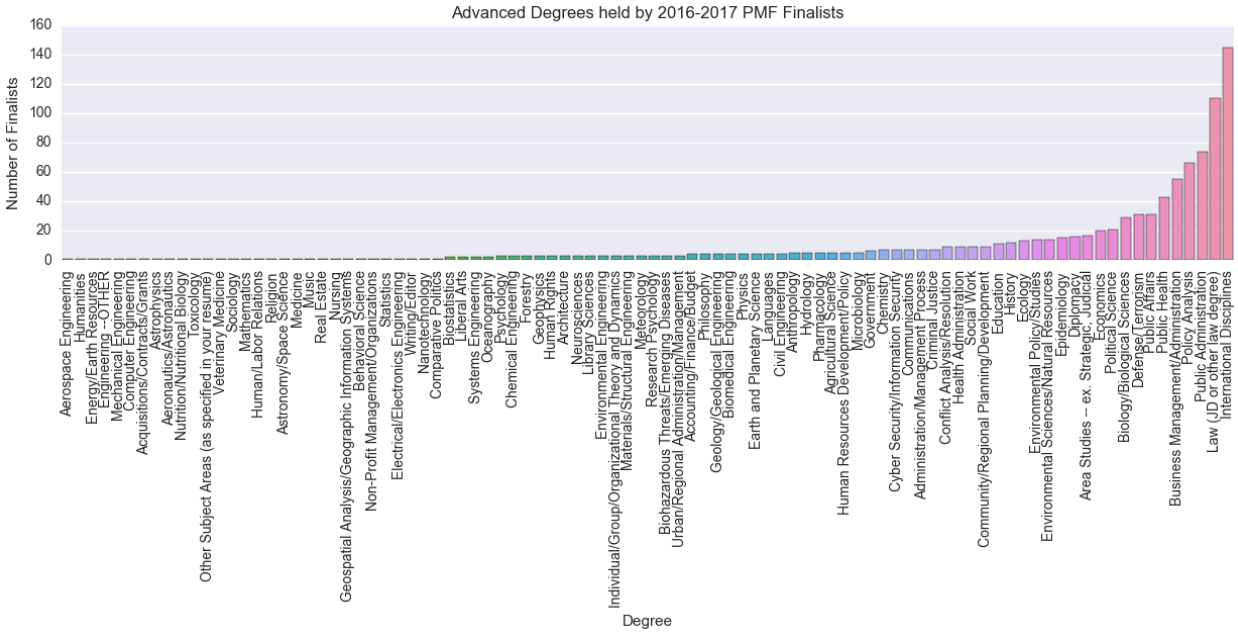


Figure 8: Count plot of advanced degrees held by the combined 2016 and 2017 classes.

Caution should be used in attempting to use this distribution to predict the makeup of future classes, because there was a fair amount of variance in the distribution of Finalists between 2016 and 2017. This stays true even after normalizing for the different class sizes. Looking just at the degrees with a relatively strong representation in both classes, many saw substantial swings in the number of PMFs with that degree between the years.

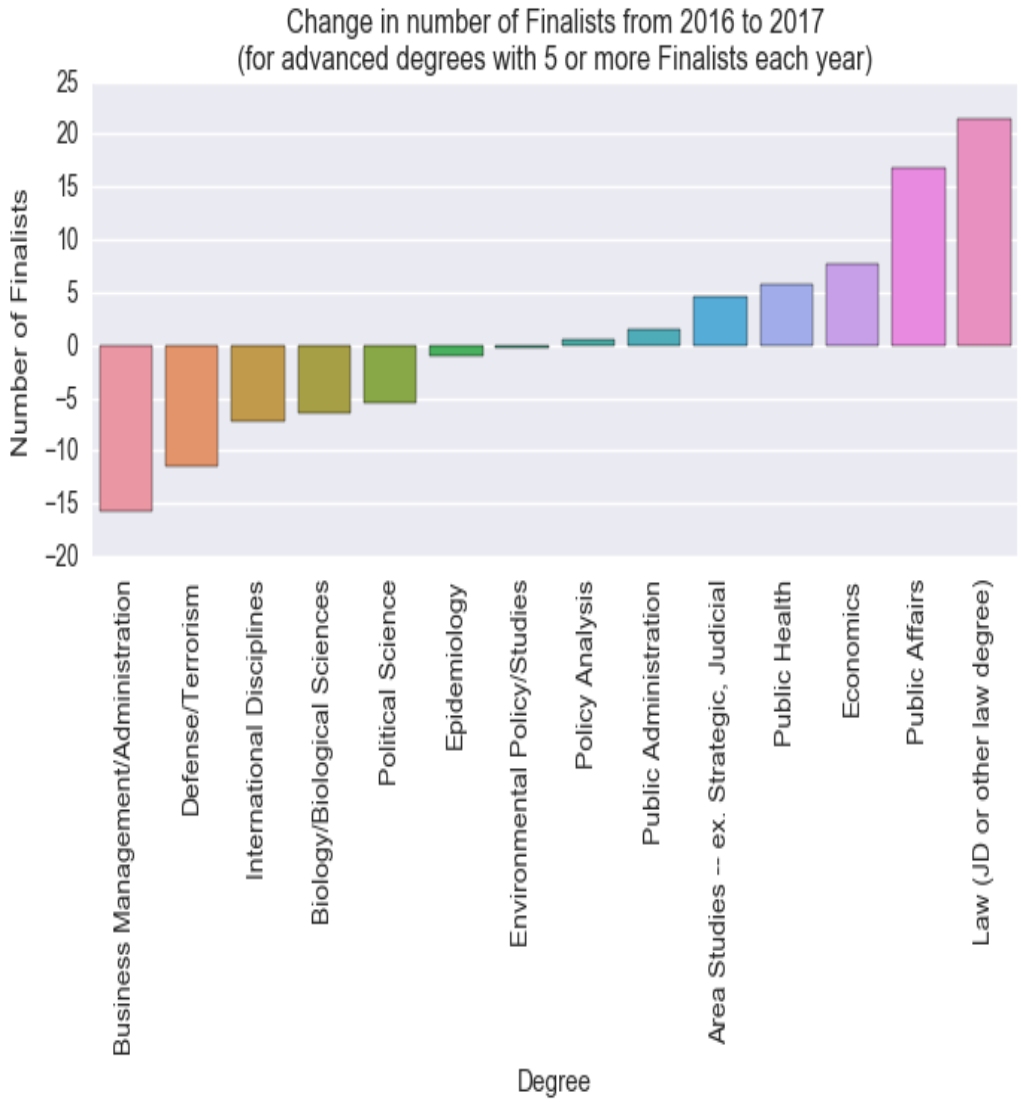


Figure 9: Changes in the number of Finalists with particular degrees admitted in 2017, relative to 2016. Numbers may not land on integers because of scaling to account for different class sizes.

For example, the 2016 class had many more Business Management/Administration degree holders while the 2017 had a larger number of lawyers. I asked the PMF office about this difference, and was told it had more to do with the different advanced degree makeup of the applicants and less with different priorities for federal hiring that year.

I'm not completely satisfied with that explanation. While the average appointment rate for the 2016 Finalist class was 57%, certain degrees had more or fewer appointed Finalists than the average

appointment rate would predict. Comparing the over/under performance of those degrees in 2016 to the change in scaled number of Finalists in 2017, there are hints of some level of correlation.

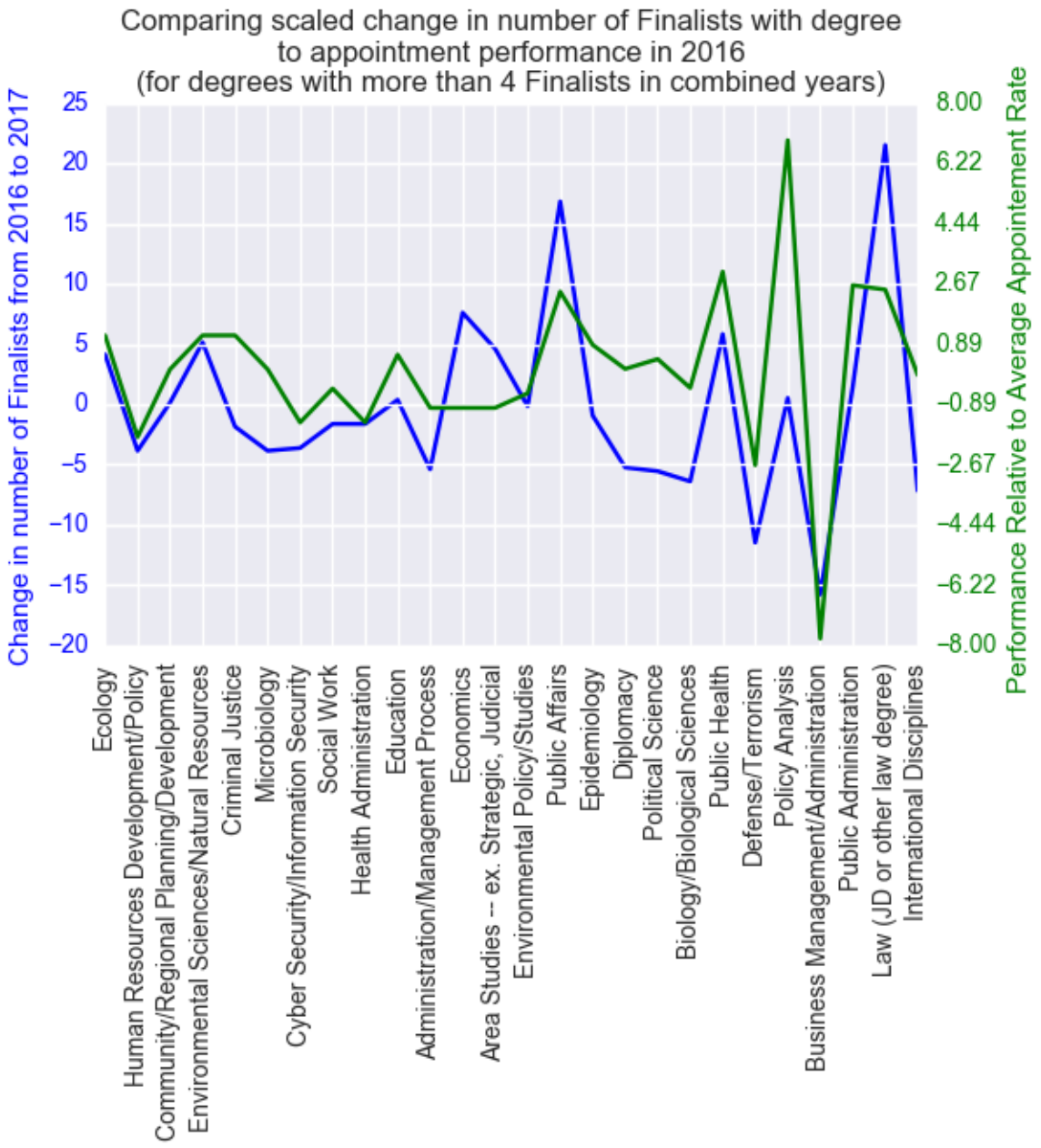


Figure 10: Comparing the over/under performance of degrees in 2016 to getting appointed with the change in relative number of Finalists admitted in 2017.

Quantifying that correlation is something I still need to do. It could be significant or it could not, but I share it here as it might be interesting for others to consider.

A csv file with the detailed information on the 2016-2017 classes (excluding names) is on GitHub, as is the python files used to collect appointment rates for the 2017 class and to do analysis of the 2016-2017 class data.

Resources	Filename	Description
1	FinalistsDetailed1617-GitHub.csv	Detailed information on Finalists from 2016-2017
2	Appointments2017-GitHub.csv	2017 Appointment rate csv
2	FinalistsDetails20162017-GitHub.ipynb	Python notebook for analysis of detailed information
3	Appointment_Rates_Github.ipynb	Python web scraper and basic plots for 2017 appointment rates

The python file FinalistsDetails20162017-GitHub contains the scripts for making these plots and additional plots covering the full set of degrees—not just the most population ones.

Trends in Available Jobs.

Finally, as a Finalist I have access to the Talent Management System (TMS), or job board. About a month after becoming a Finalist, I started recording the jobs posted on the TMS. I missed a few postings, but have records for the vast majority of jobs that have been offered, as of the writing of this report, to the 2017 PMF Finalist class.

At the beginning, I was saving the data by copy/pasting the job information into excel spreadsheets. But the deluge of postings that coincided with the virtual hiring fair convinced me to write a Python web scraper to automatically collect the job data. All but 23 of the job postings I have recorded so far were collected with the web scraper.

As of this writing, there have been 361 postings to the TMS (plus the few I missed recording at the start of being a Finalist). Many of those postings are for multiple positions (e.g. a single NIH posting was for 41 positions). The total number of potential positions in the federal government that have been posted on the TMS to date for 2017 PMFs is 920.

Number of Postings	Number of Positions	Positions per Finalist
361	920	2.2

Both those numbers are larger than reality. Several of the postings were mistakes, repeats, or postings about webinars and not job announcements. In addition, several agencies had a max number of PMFs they were interested in hiring spread over many postings, but my little python program recorded the max number as the number of positions available for each posting. Perhaps in the future I'll tackle these issues, but for now I'm treating them as minor enough to be ignored.

The 2017 Finalists were announced in early January of 2017. However, at that time there were no jobs on the job board for overly-eager Finalists to blast with applications. A few trickled in during late January and early February, but it wasn't till the virtual hiring fair on February 28th that the jobs really started to roll in. As I got my web scraper up and running, I could take recordings of the TMS (job board) every few days and use the posted start/end dates to make a timeline of the number of postings and positions available on the TMS on any given day.

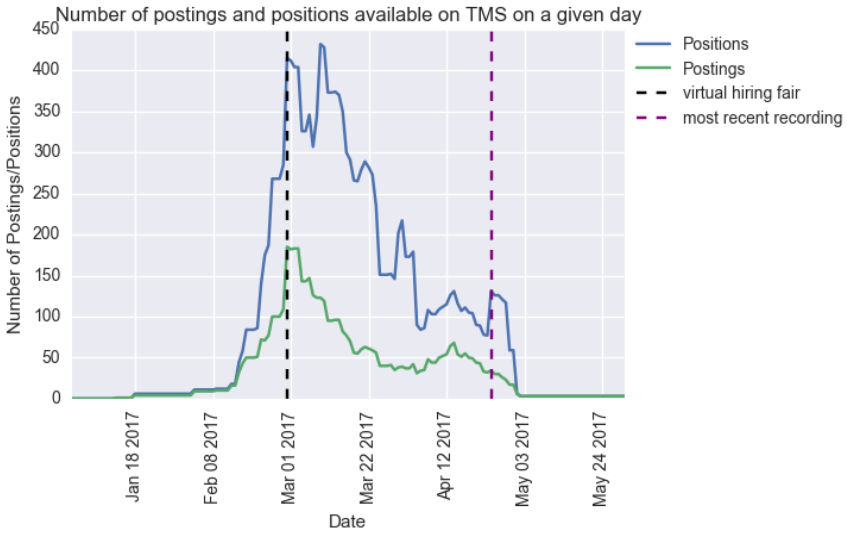


Figure 11: Plots of the number of postings and positions active on the TMS over time. Listings past the vertical purple line use the end date for currently open jobs.

Since the peak just after the job fair, there's been a slow drop-off in the number of jobs available on the job board, with occasional spikes when additional positions were posted. During a webinar, the PMF office staff stated they normally see two spikes in postings to the job board: one around the time of the job fair and a second, smaller peak at the end/start of the budget year.

Not all agencies posted jobs at the same time. While most followed the trend of posting positions near the virtual hiring fair, others were early (e.g. Dept. of Ag: NRCS) or were substantially delayed (e.g. Dept. of Health and Human Services: NIH). The new president, questions about budgets, and/or attempts to play the Finalist field are just some of the explanations that I've come up with for this, but I have no way to determine the cause of this variance.

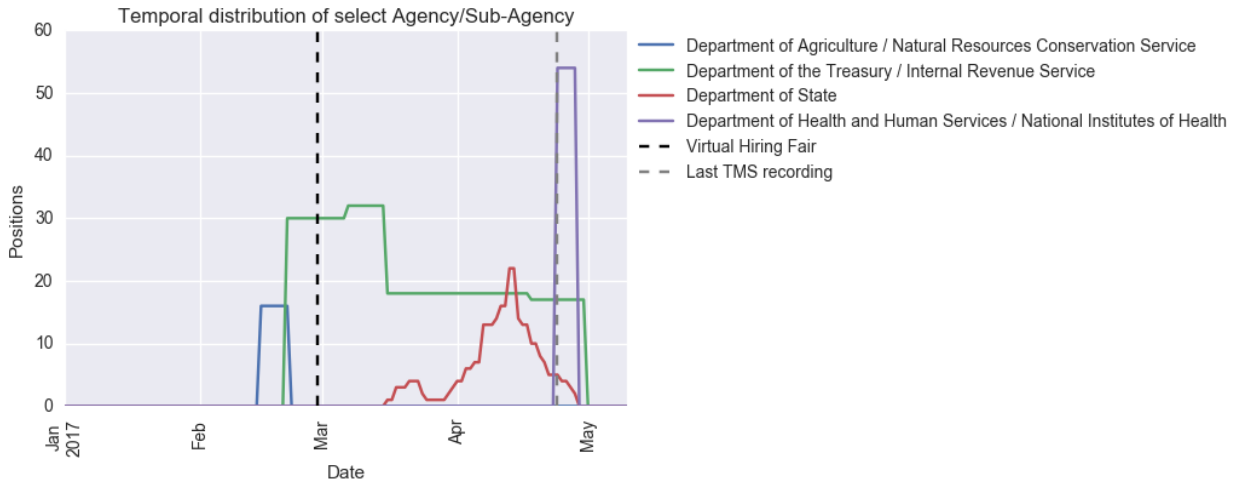


Figure 12: Timeline of when select agencies and sub-agencies posted opportunities on the TMS. Agencies chosen to highlight the differences in posting times.

There was also a substantial variation in the number of days a job was open on the job board.

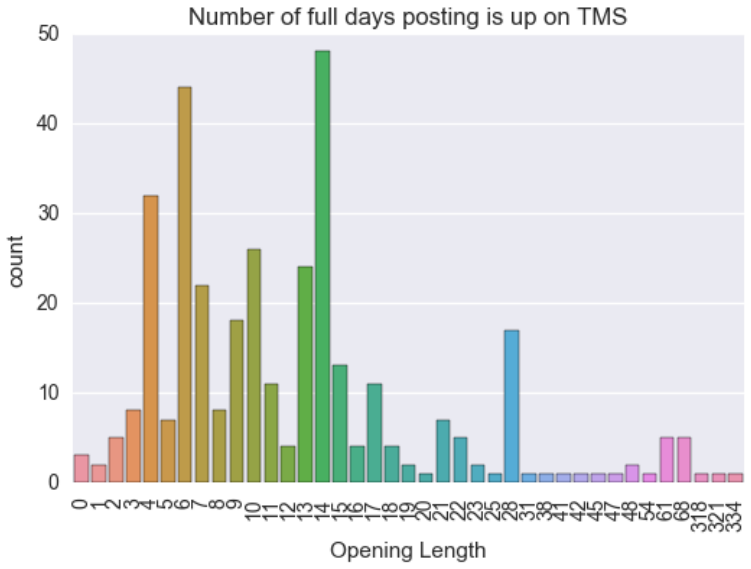


Figure 13: Plot of the number of days postings were up on the TMS.

The Broadcasting Board of Governors was not only was the first agency to post on the job board for 2017, they also hold the record for the longest open position: 334 days.

A few postings were only open for less than a day. Many of these weren't real positions, but rather webinars, mistakes, or some other bit of noise that's been added to the system.

The average posting was open for just under 2 weeks.

In years past, I believe agency hiring managers were not required to post jobs on the job board [5]. Instead, if they found a Finalist they liked, they could hire them for a position without worrying about other Finalists applying.

This year every job had to be on the job board. One hiring manager I was in contact with did things the old way (reach out to Finalists directly, conduct interviews, pick the favorite candidate), before posting a job to the job board. That hiring manager told me their HR office wanted the job up on the TMS for 5 work days before they could make a decision. Another agency realized they wanted to offer me a position at a higher grade than initially advertised, so posted another position almost exactly like the first but with a higher starting grade. That posting was only open for a few days as well.

These are just anecdotal, of course, but it's a possible explanation why many of the jobs were only open for a few days.

One of my original questions when thinking of applying to the PMF program was which agencies hire PMFs? The PMF office keeps an online, pseudo-up-to-date list of agency coordinators—the official point of contact (POC) between PMF Finalists and participating agencies [4]. Last time I checked, there were 125 names on that list, although some are duplicates and others are POC for sub-agencies within a larger body. Removing those from the count, there are 45 unique agencies in the list. And because every agency that participates in the PMF program must list a coordinator, this number gives me a good upper bound on the number of agencies that could post opportunities on the TMS.

As of the writing of this report, 27 agencies had posted at least 1 position on the job board. I have no knowledge if this is typical or a deviation from normal because of external factors (e.g. new president, budget, and a hiring freeze). Looking at which agencies are offering the majority of open positions, just 3 agencies account for about half of the total available positions:

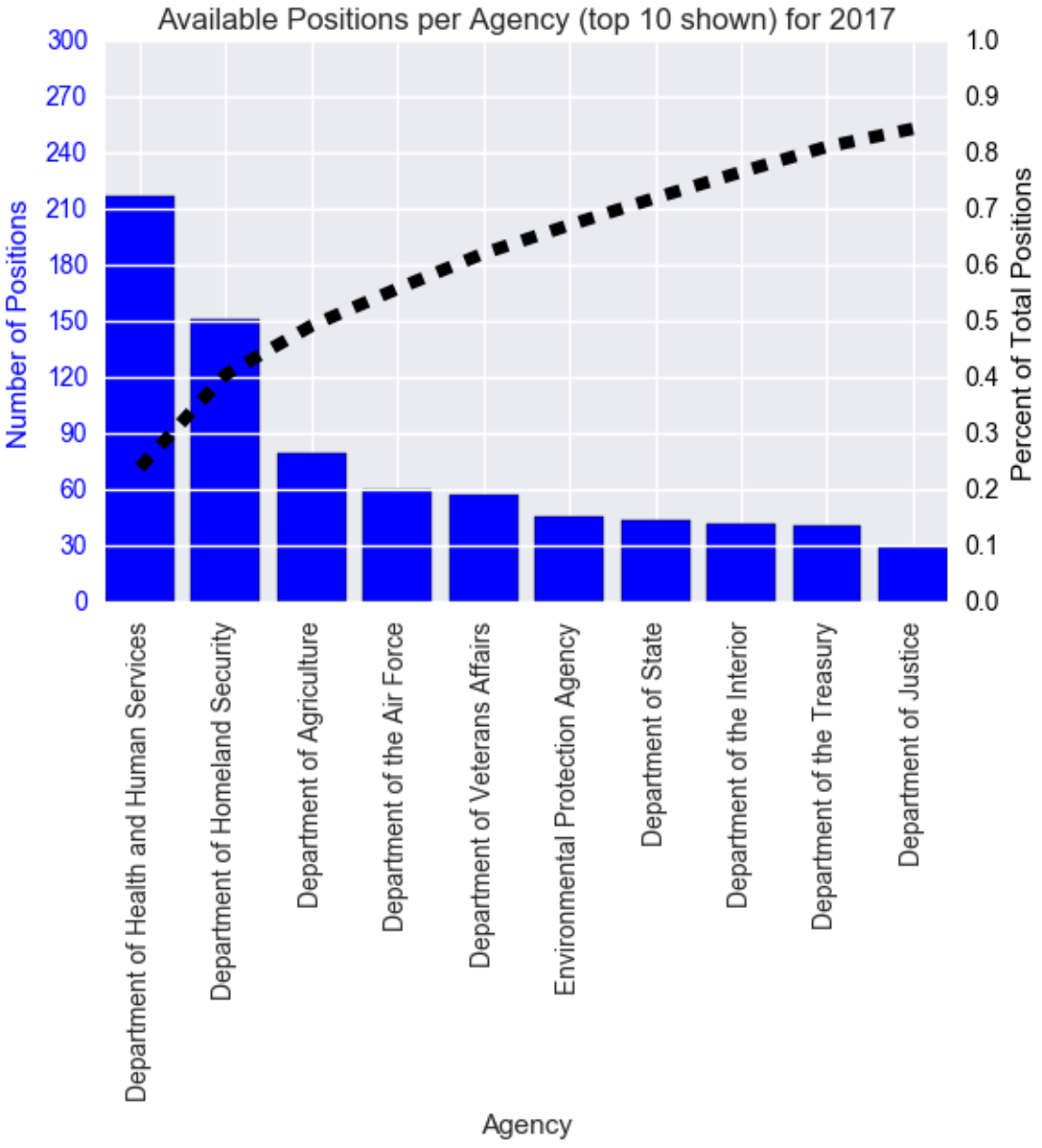


Figure 14: Number of positions per agency shown in blue, with percent of total positions contained in top N agencies shown in black.

Before applying to the PMF program, I was also interested in what types of jobs were available to PMFs. With 91 unique advanced degrees in the 2016-2017 Finalist classes and 45 partner agencies, perhaps it is not surprising that over 150 unique job titles have appeared on the TMS.

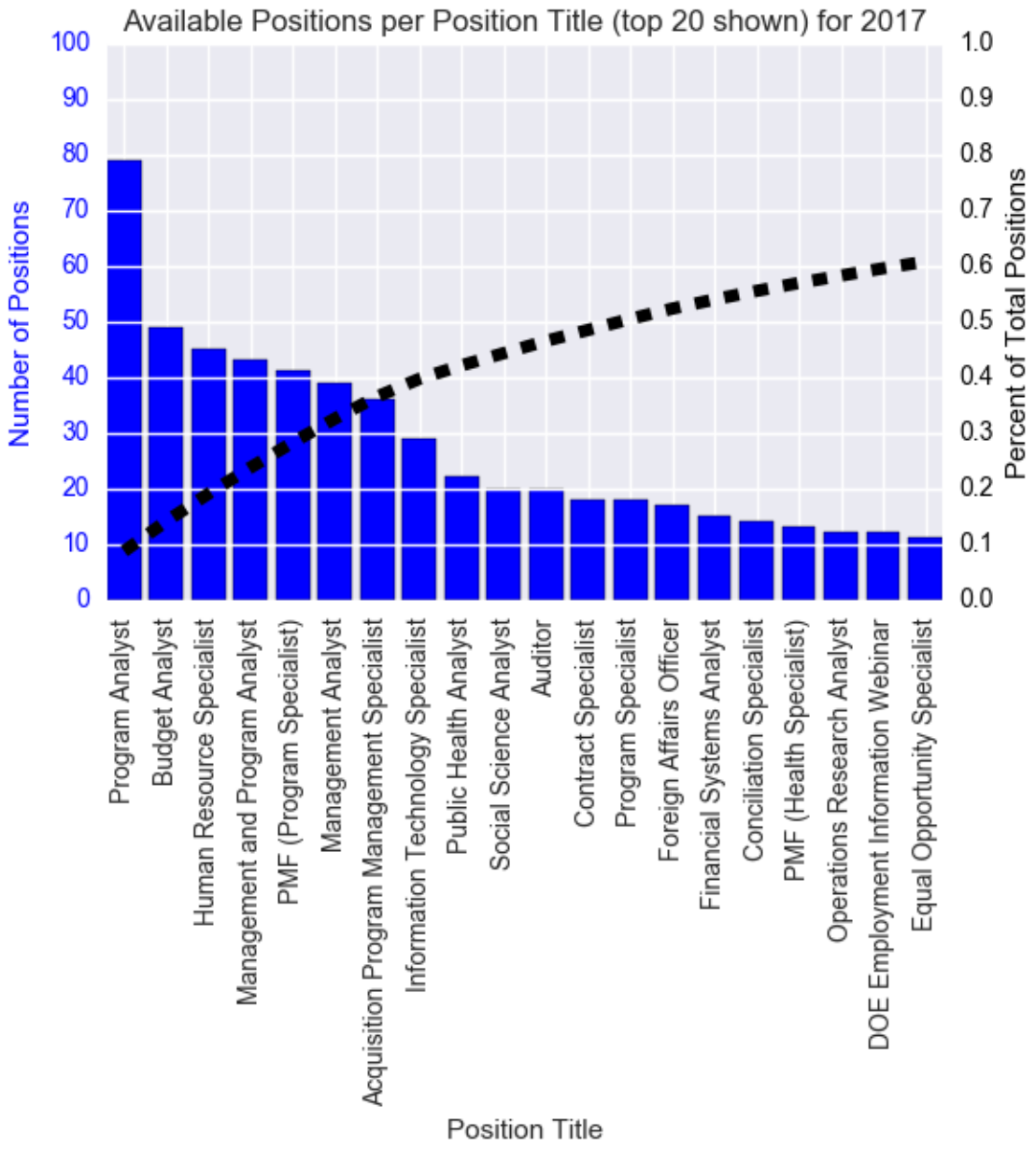
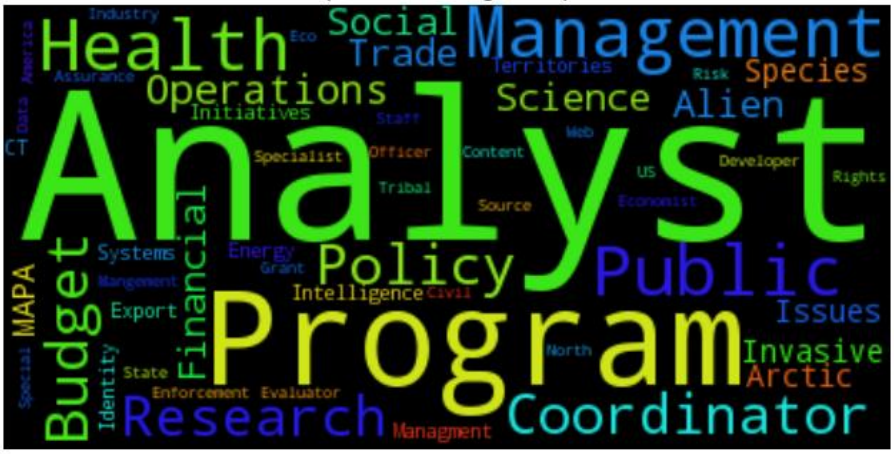


Figure 15: Number of positions per position title shown in blue, with the percent of total positions in the top N titles shown in black. Note that the agency chooses the job title and I didn't try to combine similar titles. So silly results can occur, like 'Program Specialist' and 'PMF (Program Specialist)' being counted as 2 different titles.

However, over 75% of these job titles include one of two words: Analyst or Specialist.

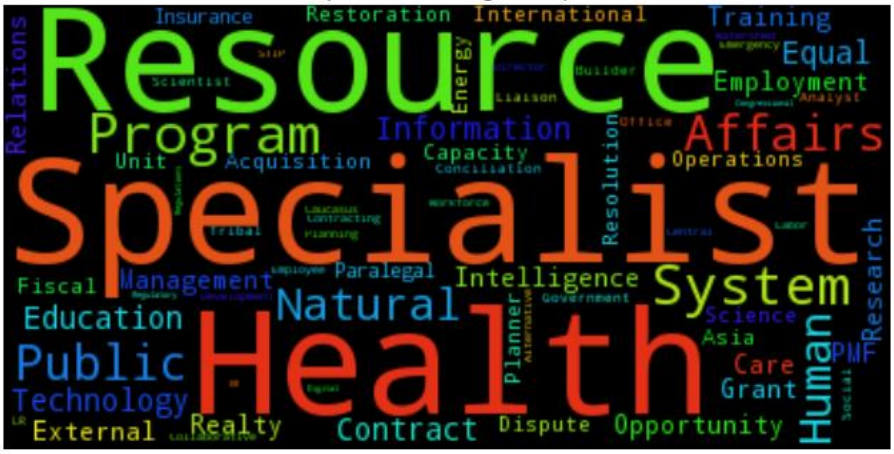
Analyst positions account for 38.9% of the positions, although there are many different types of analyst positions. Here's a word cloud of the words in Analyst position titles.

Position Title word cloud with keyword: Analyst
(words with length >=1)

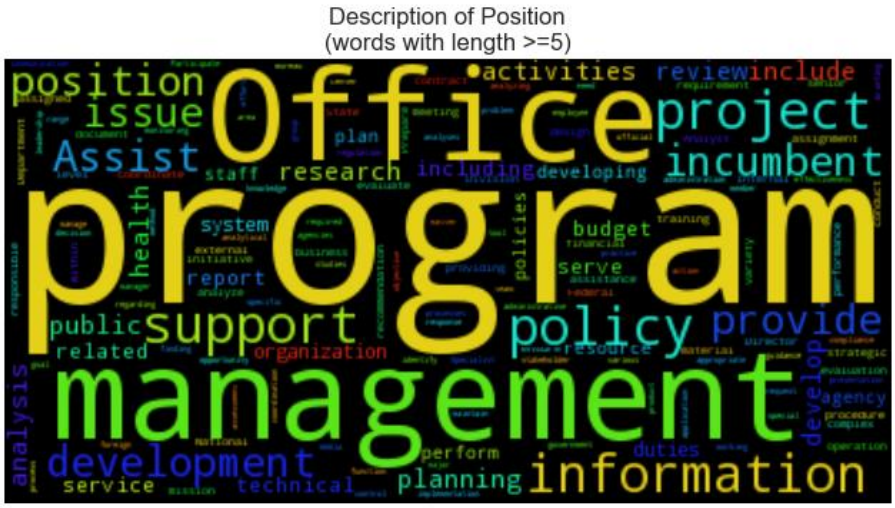


Specialist positions account for 37.5% of the positions. Here's a word cloud of the words in Specialist position titles.

Position Title word cloud with keyword: Specialist
(words with length >=1)



The words used to describe the positions might also provide some insight. Here's a word cloud built from the longer words used in the 'Description of Position' section of each posting (note this cloud only uses the postings collected with the Python web scraper):



I'm still working on a clever way to analyze this information, but perhaps the word clouds will be useful in giving future PMF applicants an idea about what's available.

Of course, around a quarter of the positions are for opportunities with job titles that include neither Analyst nor Specialist. Here's a count plot of the number of opportunities for each of these titles:

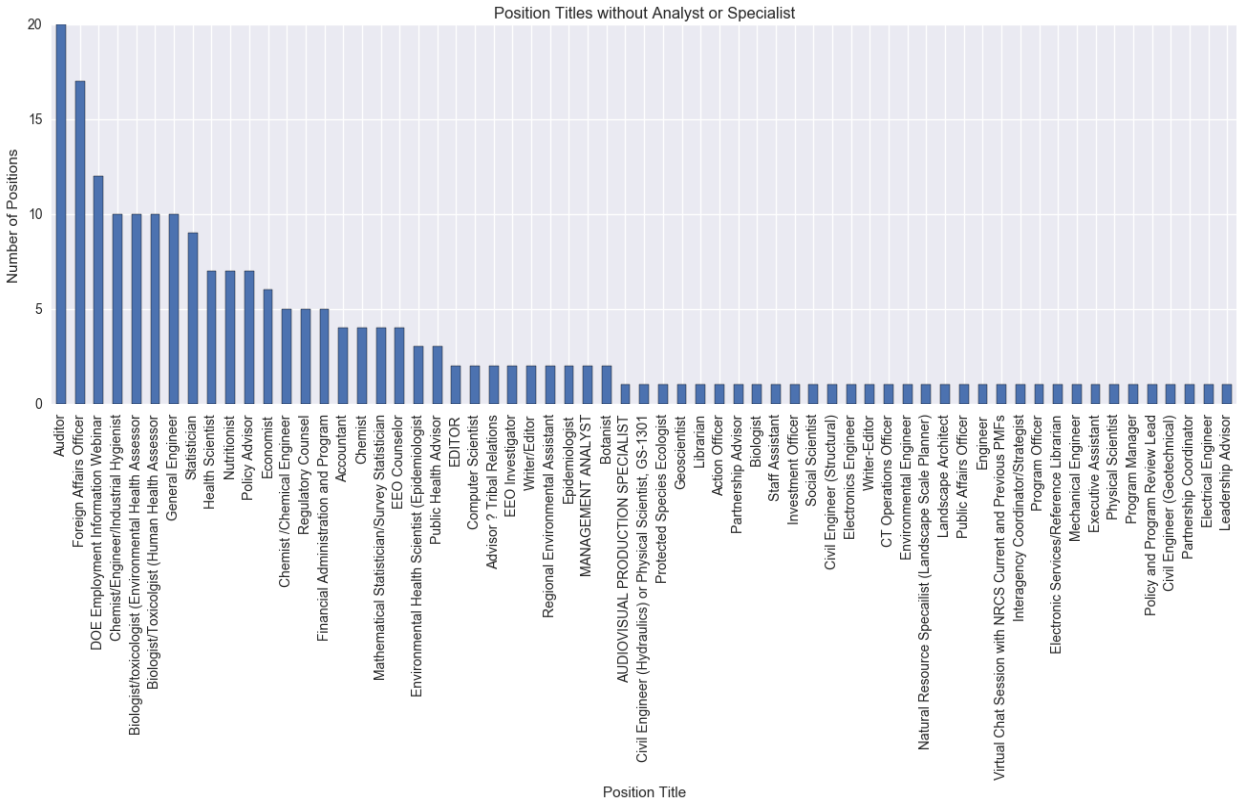


Figure 16: Number of positions for jobs that don't include the words 'Analyst' or 'Specialist' in the title. Note that this is case sensitive. Another thing I could improve at a later date.

Finally, the majority (57.5%) of positions are located in D.C. That number climbs to 77.1% if you include neighboring Maryland and Virginia.

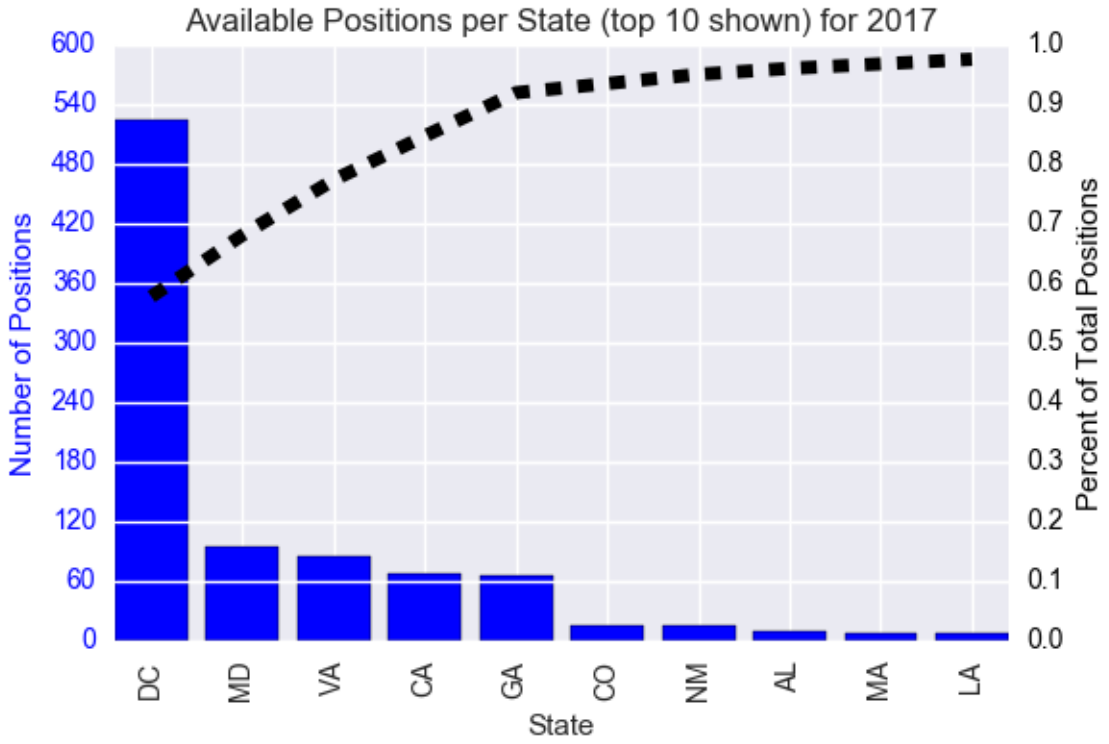


Figure 17: Number of positions per state shown in blue, with the percent of total positions contained in top n states shown in black.

A csv file of the meta data on each of the TMS postings is saved on GitHub (note: some postings are split onto multiple rows. This happens if the posting had positions in multiple locations). The web scraper used to collect the information off of the TMS, the python program used to condense all the scraped data into a single csv file, and the python program used to do the analysis are also on GitHub.

Resources	Filename	Description
1	meta_composite_GitHub.csv	Combined meta data on jobs posted to TMS for 2017
2	PMF_TMS_Scrape_and_Study_Github.ipynb	Python notebook for scraping information off of TMS
3	Combine Meta Data-GitHub.ipynb	Python notebook to combine csv files recorded from web scraper into single file
4	Combined_Analysis_GitHub.ipynb	Python notebook for analysis of TMS data

Conclusion

In this report, I tried to combine the most interesting work I've done on the PMF program since I first learned of its existence. The three main sections of this report consider how Finalists selection has changed from 2012-2017, the distribution of degrees and appointments in the recent Finalist classes, and the positions and agencies that have been available to the 2017 class.

This is an ongoing project, and I hope to periodically update this report as new data is gathered, new methods for analysis are developed, and as new questions arise. I'm very curious to see if the streamlined PMF application process introduced in 2017 might began to counteract the trends in application number and pipeline schools observed in this report. I'm also impressed by the number of potential positions open to PMFs in the federal government. Because the appointment rate for PMF Finalists has been around 50-60%, I had assumed there would not be enough positions to go around. However, that doesn't seem to be the case. I now believe it's most likely that a good portion of the 40-50% of Finalists that don't accept positions do so because none of the opportunities appeal to them.

There are many limitations on conclusions which can be drawn from this work. I've only recorded a few months of data for the 2017 class, only been able to look at degree and appointment distributions for 2 classes, and only have used records dating back to 2012 in comparing admission rates. Furthermore, the changing political and economic environment, plus future changes to the PMF program, could easily alter any trends found in this report.

With those caveats aside, I hope this report was interesting and/or useful. As mentioned in the introduction, selected raw data, figures, and code may be found at https://github.com/keklarup/PMF.

[1] "PMF as a Pathway," 20 04 2017. [Online]. Available: https://www.pmf.gov/opportunity/index.aspx.

[2] "PMF-STEM," pmf.gov, [Online]. Available: https://www.pmf.gov/the-opportunity/pmf-stem.aspx. [Accessed 20 04 2017].

[3] "Kullback–Leibler divergence," Wikipedia, [Online]. Available: https://en.wikipedia.org/wiki/Kullback%E2%80%93Leibler_divergence. [Accessed 22 04 2017].

[4] "Become a PMF FAQ," pmf.gov, [Online]. Available: https://pmf.gov/become-a-pmf/faqs.aspx#q12. [Accessed 22 04 17].

[5] "Agencies: Find Candidates," pmf.gov, [Online]. Available: https://pmf.gov/agencies/find-candidates.aspx. [Accessed 25 04 2017].

[6] "PMF Coordinators," pmf.gov, [Online]. Available: https://apply.pmf.gov/coordinators.aspx. [Accessed 14 01 2017].